§12.8 Lagrange multipliers and Sensitivity

Let $\bar{x}$ be a local min to (NLP) $\begin{cases} \min f(x) \\ \text{s.t. } c_i(x) = 0 & i = 1, \dots, m_E \\ \quad c_i(x) \geq 0 & i = m_{E+1}, \dots, m \end{cases}$

then 1st order nec. cond $\Rightarrow$ $\exists \bar{\lambda}$ s.t. $\nabla_x L(\bar{x}, \bar{\lambda}) = 0$

where $L(x, \lambda) = f(x) + \sum_{i=1}^{m} c_i(x) \lambda_i$.

is the Lagrangian.

$\bar{\lambda}$ indicates sensitivity of obj function to constraint $c_i(\bar{x})$.

To illustrate this

- When $i \notin \mathcal{A}(\bar{x})$ we have $c_i(\bar{x}) > 0$ and $\underline{\bar{\lambda}_i = 0}$ (complementarity)

  $\rightsquigarrow$ function value at $\bar{x}$ is indep of having this constraint.

  If we change constraint slightly $\rightarrow$ $\tilde{c}_i(\bar{x})$ remains inactive (>0) and $\bar{x}$ remains a local min

- When $i \in \mathcal{A}(\bar{x})$, $c_i(x) \geq 0$. is st. $c_i(\bar{x}) = 0$.

  If we change constraint slightly, carefully chosen to use Taylor

  $$c_i(x) \geq -\varepsilon \|\nabla c_i(\bar{x})\|$$

  with ε sufficiently small s.t. $\bar{x}(\varepsilon)$ = perturbed sol.

  $$\mathcal{A}(\bar{x}(\varepsilon)) = \mathcal{A}(\bar{x})$$

  and $\bar{\lambda}(\varepsilon) \approx \bar{\lambda}$

$\Rightarrow$ $c_i(\bar{x}(\varepsilon)) = -\varepsilon \|\nabla c_i(\bar{x})\|$ (active constraint)

$\Rightarrow$ $c_i(\bar{x}(\varepsilon)) - \underbrace{c_i(\bar{x})}_{=0} = -\varepsilon \|\nabla c_i(x)\|$

Taylor SS

$$\nabla c_i(\bar{x})^T (\bar{x}(\varepsilon) - \bar{x})$$

For all other active constraints $j \in \mathcal{A}(\bar{x})$, $i \neq j$:

$$0 = \underbrace{c_j(\bar{x}(\epsilon))}_{=0} - \underbrace{c_j(\bar{x})}_{=0} \approx \nabla c_j(\bar{x})^T(\bar{x}(\varepsilon) - \bar{x})$$

$$\Rightarrow f(\bar{x}(\varepsilon)) - f(\bar{x}) \overset{\text{Taylor}}{\simeq} \nabla f(\bar{x})^T (\bar{x}(\varepsilon) - \bar{x})$$

$$\doteq \Big(-\sum_{j \in \mathcal{A}(\bar{x})} \bar{\lambda}_j \nabla c_j(\bar{x})\Big)^T (\bar{x}(\varepsilon) - \bar{x})$$

$$\approx -\varepsilon \|\nabla c_i(\bar{x})\| \lambda_i$$

$$\Rightarrow \frac{d f(\bar{x}(\varepsilon))}{d\varepsilon} = -\|\nabla c_i(\bar{x})\| \lambda_i$$

§13 Linear Programming

Let $c \in \mathbb{R}^n$, $b \in \mathbb{R}^m$, $A \in \mathbb{R}^{m \times n}$

The problem

$$\boxed{\begin{cases} \min & c^T x \\ & Ax = b \\ & x \geq 0 \end{cases}}$$

is a linear program (LP) in standard form

Other LPs can be transformed in standard form.

Take for example constraints $Ax \geq b$.

We introduce variables (≥ 0) x_+, x_-, s ← slack variables

then if we let $x = x_+ - x_-$

$$Ax \geq b \iff Ax_+ - Ax_- = b + s$$

$$\iff (A, -A, -I)\begin{pmatrix} x_+ \\ x_- \\ s \end{pmatrix} = b$$

$$\begin{pmatrix} x_+ \\ x_- \\ s \end{pmatrix} \geq 0.$$

Assumptions: $A \in \mathbb{R}^{m \times n}$, $m < n$, $\text{rank}(A) = m$.

If $m \geq n$ then most likely $\mathcal{F} = \emptyset$ or only one sl
if $\text{rank}(A) < m$, redundant eq can be identified using e.g. QR decomp. and then removed.

Note: LP is a convex problem (feasible set & obj fun are convex)
$\Rightarrow$ All local minima are global

Note: Regularity of LP.

$$\mathcal{F} = \{x \in \mathbb{R}^n \mid Ax = b, x \geq 0\}$$

$$T(\mathcal{F}, \bar{x}) = \{d \in \mathbb{R}^n \mid Ad = 0, d_i \geq 0, i \in \{1..n \mid \bar{x}_i = 0\}\}$$

$$= L(\mathcal{F}, \bar{x})$$

Optimality Conditions

If $\bar{x}$ solves (LP) then there exists $\bar{y} \in \mathbb{R}^n$ & $\bar{z} \in \mathbb{R}^m$ s.t.

$$\boxed{(KKT) \begin{cases} \bar{z} \geq 0 \\ \bar{z}^T \bar{x} = 0 \\ c - A^T \bar{y} - \bar{z} = 0 \\ A\bar{x} = b, \bar{x} \geq 0 \end{cases}}$$

positivity
complementarity
$\nabla_x L(\bar{x}, \bar{y}, \bar{z}) = 0$
feasibility

check: $\underbrace{L}_{\substack{\text{Lagrange mult.} \\ = \lambda}}(x, y, z) = c^T x - y^T(Ax - b) - z^T x$

(Same sign convention as in Nocedal & Wright now)

Necessary optimality conditions are also sufficient for LP:

If $\bar{x}, \bar{y}, \bar{z}$ solves KKT then

$$c^T \bar{x} = (A^T \bar{y} + \bar{z})^T \bar{x}$$

$$= \bar{y}^T \underbrace{A\bar{x}}_{= b} + \underbrace{\bar{z}^T \bar{x}}_{\substack{= 0 \\ \text{(compl)}}} = b^T \bar{y}$$

For any other feasible point x:

let $d = x - \bar{x}$ then

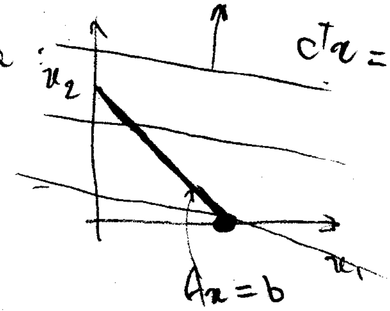
$$Ad = 0, \quad d_i \geq 0 \text{ for } \bar{x}_i = 0$$

Thus
$$\begin{aligned} c^T x &= c^T(\bar{x} + d) \\ &= (A^T\bar{y} + \bar{z})^T(\bar{x} + d) \\ &= b^T\bar{y} + \bar{y}^T \underbrace{Ad}_{=0} + \bar{z}^T d, \end{aligned}$$
where $z_i d_i$ is $\begin{cases} \geq 0 & \text{if } \bar{x}_i = 0 \\ = 0 & \text{if } \bar{x}_i > 0 \end{cases}$

$\Rightarrow \quad c^T x \geq b^T \bar{y} = c^T \bar{x}$

$\Rightarrow \quad \bar{x}$ is a min of (LP).

2D example:



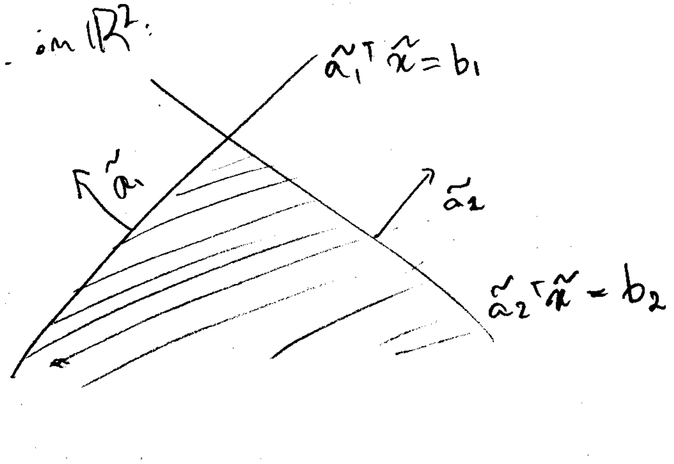
$c^T x = v$ (set of points w/ same obj. fun value)

min in this case is a vertex.

<u>Def</u>: (Polyhedron) The intersection of finitely many halfspaces
$$\{\tilde{x} \in \mathbb{R}^n \mid \tilde{a}_i^T \tilde{x} \leq b_i, \ i = 1, \dots, \tilde{m}\}$$
is called a <u>Polyhedron</u>.

A <u>bounded polyhedron</u> is a <u>polytope</u>

e.g. in $\mathbb{R}^2$:



<u>note</u> $\tilde{a}_i$ point <u>OUT</u> of polyhedron
(sign of inner prod)

- Polytopes are convex ($\cap$ of convex sets is convex)

Any set of the form

$$\{\tilde{x} \in \mathbb{R}^{\tilde{n}} \mid \tilde{a}_i^T \tilde{x} \leq \tilde{b}_i, \ i = 1..m\}$$

Can be described by:

$$\{x \in \mathbb{R}^n \mid Ax = b, \ x \geq 0\}, \quad n = 2\tilde{n} + m$$

($2\tilde{n}$: $\tilde{x}_+, \tilde{x}_-$; m: slacks)

This is a polyhedron in <u>standard form</u>.

<u>Def</u>: Let $\mathcal{F} = \{x \in \mathbb{R}^n \mid Ax = b, x \geq 0\}$ be a polyhedron in standard form.

A vector $\bar{x} \in \mathcal{F}$ is a <u>vertex of</u> of $\mathcal{F}$ if:

$$\bar{x} = \lambda x_1 + (1-\lambda) x_2, \quad x_1, x_2 \in \mathcal{F}, \quad \lambda \in (0,1)$$

$$\Rightarrow \quad x_1 = x_2$$

vertices are a.k.a <u>extreme points</u>

vertices are points that cannot be written as a strict convex combination of points in the polyhedron.

<u>Dual problem</u>

$$(DLP) \begin{cases} \max & b^T y \\ \text{s.t} & A^T y \leq c \end{cases}$$

is also a linear problem (regularity at every pt guaranteed)

$(A \in \mathbb{R}^{m \times n}, \ b \in \mathbb{R}^m, \ c \in \mathbb{R}^n)$

Optimality conditions for (DLP):

$\bar{y}$ solves DLP iff $\exists \bar{x} \in \mathbb{R}^n$ s.t.

$$\begin{cases} \bar{x} \geq 0 \\ \bar{x}^T (A^T \bar{y} - c) = 0 \\ A^T \bar{y} \leq c \\ b - A\bar{x} = 0 \end{cases}$$

Let $\bar{z} = c - A^T\bar{y}$. then clearly $\bar{z} \geq 0$

hence: $c - A^T\bar{y} - \underbrace{\bar{z}}_{\text{slacks}} = 0$ ($\nabla_z L = 0$ for (LP))

Let x be feasible for (LP) and y feasible for (DLP) then:

$$\boxed{b^T y = \underbrace{x^T}_{\geq 0} \underbrace{A^T y}_{\leq c} \leq x^T c}$$ (Weak duality

Theorem: i) If either (LP) or (DLP) has a solution then so does the other, and at the solution the objective values are equal

ii) If x and y are feasible points for (LP) and (DLP) resp. then $b^T y \leq c^T x$

iii) If either problem (LP) or (DLP) has an unbounded objective, then the other problem has no feasible point.

Example of iii)

$$(LP) \begin{cases} \min\ (-1, 0)\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \\ \text{s.t.} \\ (0 \quad 1)\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 0 \\ x_1, x_2 \geq 0 \end{cases}$$

is unbounded below

$$(DLP) \begin{cases} \max\ 0^T y \\ \begin{pmatrix} 0 \\ 1 \end{pmatrix} y \leq \begin{pmatrix} -1 \\ 0 \end{pmatrix} \end{cases}$$

no feasible points since $0 > -1$